Title: AI-Driven Prediction of Phage Receptor-Binding Proteins for Phage Therapy

Yueyi Cai[1], Xinqi Li[2], Hanshuo Lu[2], Rory John Garland[2], Faiz Mohammad Khan[2], Heather E. Allison[1]
1. Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool, Liverpool, UK
2. M&F tech, Liverpool, UK

Background:
With the global rise of antimicrobial resistance, bacteriophages have gained renewed attention as a viable alternative or complement to traditional antibiotics. A phage's receptor-binding protein (RBP) is crucial for its docking onto and infecting specific bacterial hosts. This RBP-mediated attachment underpins the success of phage therapy, allowing high specificity and minimizing unintended effects on beneficial microbiota. Efficient identification and characterization of RBPs are therefore key to designing robust phage therapies, especially when used alongside existing antibiotics.

Aim:
This study aims to apply advanced machine learning techniques to predict and validate phage RBPs, thereby streamlining the design of precision phage therapies as an adjunct to conventional antibiotic treatments.

Methods:
We developed a CatBoost model for RBP detection, incorporating ensembled features from multiple ESM2 embeddings. Novel elements included hyperparameter optimization to fine-tune performance and feature caching to handle computationally intensive transformations. A curated database of phage–bacterium systems was used for training and evaluation. We further integrated structural modelling and protein–protein interaction analysis to predict binding affinity and specificity. Predicted RBP candidates underwent preliminary in vitro validation via bacterial culture assays to confirm binding specificity and bactericidal activity.

Results:
Our approach demonstrated improvements over existing methods reported in the literature, achieving high accuracy and precision in identifying potential RBPs. Experimental validation showed strong correlation between computational predictions and observed host specificity.

Conclusion:
These findings highlight the potential of AI-driven RBP identification for next-generation phage therapeutics. By refining predictive models and expanding validation efforts, we can accelerate the development of phage-based strategies that complement antibiotic treatments, offering a promising solution to the global antimicrobial resistance challenge.